

Bogdan Suchecki, Artur Gajdos

Uniwersytet Łódzki

SYSTEM ANALIZ I PROGNOZ (SAP) – NARZĘDZIE WIZUALIZACJI I PRZETWARZANIA DANYCH

1. Wstęp

Informacja stanowi obecnie podstawowy element walki konkurencyjnej. Dzisiaj nie surowce, technologie, a nawet nie kapitał ludzki, których przepływ jest w zasadzie nieograniczony w gospodarce globalnej, lecz informacja determinuje konkurencyjność gospodarek na światowym rynku. Informacja jest takim rodzajem zasobów, który zwiększa naszą wiedzę o nas i o otaczającym nas świecie. Do głównych cech informacji należą: niezależność od obserwatora, niewyczerpalność, różnorodność, możliwość powielania w czasie i przestrzeni, możliwość przetwarzania bez zużycia. Informacje są towarem, umiejętność zaś ich gromadzenia, przetwarzania i wykorzystywania decyduje w coraz większym stopniu o miejscu jednostki w społeczeństwie i – społeczeństwa w świecie.

System informacyjny stanowi natomiast wielopoziomą strukturę pozwalającą użytkownikowi na transformowanie określonych informacji wejścia na pożądane informacje wyjścia za pomocą odpowiednich procedur i modeli. Elementami systemu informacyjnego są podmioty, informacje, technika, relacje. Jeżeli jeden z wymienionych elementów dotyczy technologii informatycznej, mówimy wtedy o informatycznym systemie informacyjnym (por. [Gajdos 2002, s. 192-199]).

Podstawowe funkcje systemów informacyjnych skupiają się na pozyskiwaniu, gromadzeniu, utrzymywaniu, aktualizacji, przetwarzaniu i przesyłaniu informacji. Informatyczne systemy informacyjne przyjmują, w zależności od pełnionej funkcji, postać prostych systemów transakcyjnych (ewidencyjno-kontrolnych), poprzez systemy wspomagania decyzji na różnym szczeblu do systemów rzeczywistości wirtualnej (por. [Gajdos 2002, s. 185-202]).

Systemy wspomaganie decyzji, poza podstawową funkcją ewidencyjną, mają szeroko rozbudowane narzędzia do analizy danych, najczęściej automatyzujące procesy obliczeniowe (optymalizacyjne, analityczne, statystyczne, prognostyczne).

System Analiz i Prognoz (SAP), autorstwa zespołu pracowników Katedry Ekonometrii Przestrzennej Uniwersytetu Łódzkiego, stanowi efekt kilkuletnich prac nad konstrukcją informatycznego systemu wspomaganie procesów prognostycznych w ramach rynku pracy – Systemu Prognozowania Popytu na Pracę (SPPP), który został wykonany na zamówienie Rządowego Centrum Studiów Strategicznych.

SAP stanowi połączenie bazy danych, programu obliczeniowego oraz technologii WWW. System umożliwia zdalne wykonywanie obliczeń, wizualizację i prezentację danych na stronach internetowych (por. [Gajdos, Szablowski 2004]). Pozwala też na zastosowanie szerokiego wachlarza technik i modeli analityczno-prognostycznych oraz zapewnia (por. [Sucheck 1998, s. 244]):

- 1) wybór pomiędzy różnymi instrumentami prognozowania,
- 2) zastosowanie mechanizmu kontroli tych instrumentów,
- 3) możliwość dostępu i działań (aktualizacji, rozszerzania, przetwarzania) na bazie danych,
- 4) analizę błędów prognozowania,
- 5) możliwość zapamiętywania wygenerowanych prognoz,
- 6) elastyczność dostosowania do konkretnych, specyficznych zadań w zależności od potrzeb.

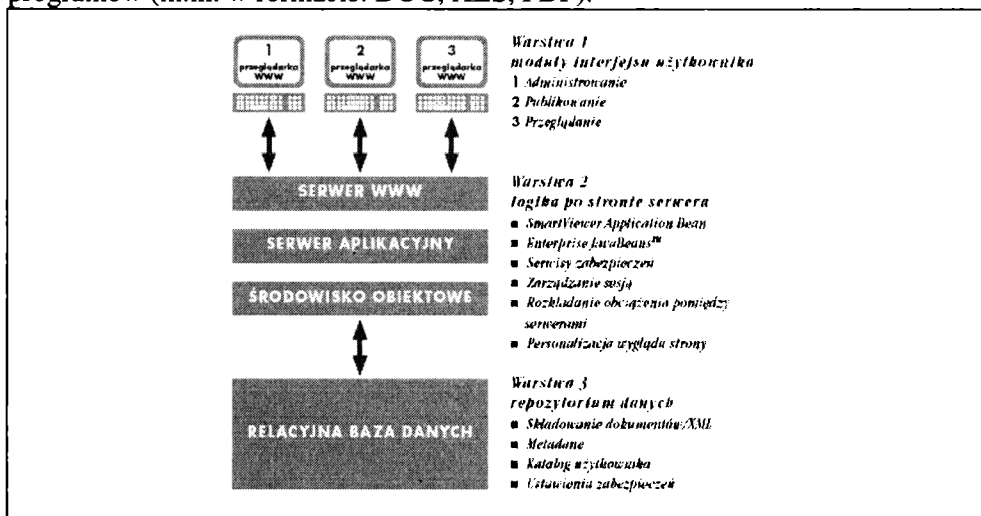
SAP pozwala, za pomocą jedynie przeglądarki internetowej, na budowanie, weryfikowanie i prognozowanie na podstawie różnych klas modeli. Począwszy od modeli analizy szeregów czasowych (trend, sezonowość, wygładzanie wykładnicze, autoregresja, wektorowa autoregresja) poprzez jednorównaniowe modele przyczynowo-skutkowe do symulacji na wielorównaniowych modelach ekonometrycznych. Generowane prognozy mogą być zapamiętywane oraz mogą stanowić podstawę dalszych analiz.

2. Informatyczne systemy analizy danych

Skomplikowanie procedur obliczeniowych oraz wykorzystywanie dużych banków danych uniemożliwiają i będą uniemożliwiać w przyszłości profesjonalną analizę danych bez wspomaganie informatycznego. Komputer nie jest jedynie narzędziem magazynowania danych, lecz spełnia (poprzez odpowiednie oprogramowanie) również funkcję analityczną oraz funkcję udostępniania danych (także zdalnie). Wiele popularnych pakietów analitycznych (Statistica, SAS, SPSS) rozwija się w kierunku umożliwiania operacji na dużych bankach danych oraz pracy zdalnej przy wykorzystaniu technologii internetowej.

SmartViewer Web Server¹

SmartViewer Web Server to serwer raportów analitycznych, służący do zarządzania wynikami analiz danych wykonanych za pomocą narzędzi SPSS. Wykorzystując SmartViewer Web Server można publikować analizy w bazie danych, do której dostęp mogą mieć, na wybranych poziomach, różni użytkownicy. Udostępniane przez SmartViewer Web Server raporty są dynamicznie tworzonymi stronami HTML, lecz posiadają funkcjonalność standardowych raportów SPSS. Uprawnieni użytkownicy mogą w przystępny sposób korzystać z raportów analitycznych, używając jedynie przeglądarki internetowej (bez konieczności instalacji dodatków) oraz nie muszą korzystać ze specjalnego oprogramowania analitycznego. SmartViewer Web Server magazynuje centralnie wszelkie analizy w postaci tabel, wykresów, raportów OLAP, dokumentów HTML oraz publikacji innych programów (m.in. w formacie: DOC, XLS, PDF).



Rys. 1. Schemat działania SmartViewer Web Server

Źródło: www.spss.com.

Analityk opracowuje i udostępnia raporty, korzystając z programu SPSS. Następnie ich odbiorcy uzyskują dostęp poprzez SmartViewer Web Server, korzystając ze standardowej przeglądarki internetowej. Informacje umieszczone w raportach mogą być swobodnie przeglądane, a dzięki interaktywnej postaci możliwa jest dalsza analiza.

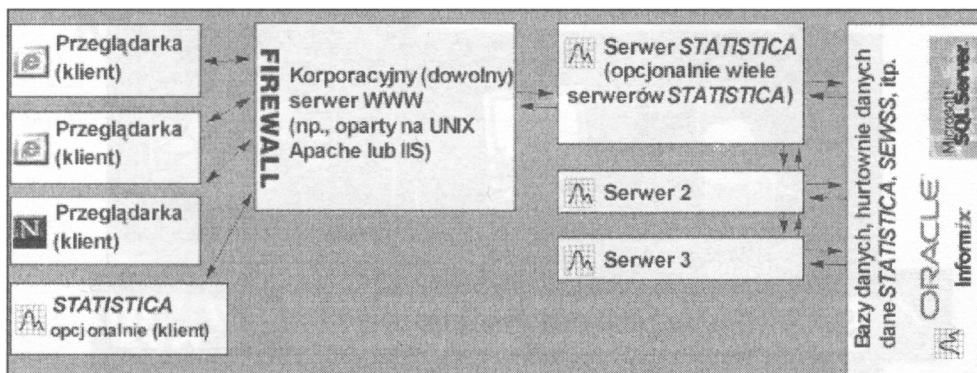
Web Statistica Server²

Web Statistica Server umożliwia wykorzystanie internetu podczas pracy z programami z rodziny Statistica. W szczególności pozwala na zdalne uruchamianie

¹ Opracowano na podstawie: www.spss.com.

² Opracowano na podstawie: www.statsoft.pl.

Statistica w oknie przeglądarki internetowej, a także szybki i łatwy dostęp do danych i narzędzi analitycznych z praktycznie każdego komputera połączanego z internetem. Web Statistica Server zapewnia funkcjonalność narzędzi analitycznych, graficznych i narzędzi do budowy zapytań i raportów programu Statistica poprzez prosty w użyciu, interaktywny interfejs przeglądarki internetowej.



Rys. 2. Schemat działania Web Statistica Server

Źródło: www.statsoft.pl.

Jest to w pełni skalowalny, bazujący na technologii internetowej system gromadzenia i analizy danych. System został zbudowany z myślą o technologii przepływu informacji i w pełni wykorzystuje wielowarstwową architekturę klient–serwer, gdzie:

warstwa 1 – jest to interfejs użytkownika na komputerze klienta (zwykła przeglądarka),

warstwa 2 – jest to oprogramowanie Web Statistica Server i implementacja *business intelligence*,

warstwa 3 – są to bazy danych lub hurtownie danych.

W przypadku tej wersji Statistica wszystkie obliczenia odbywają się na serwerze, a stacja robocza użytkownika obsługuje jedynie jego interfejs.

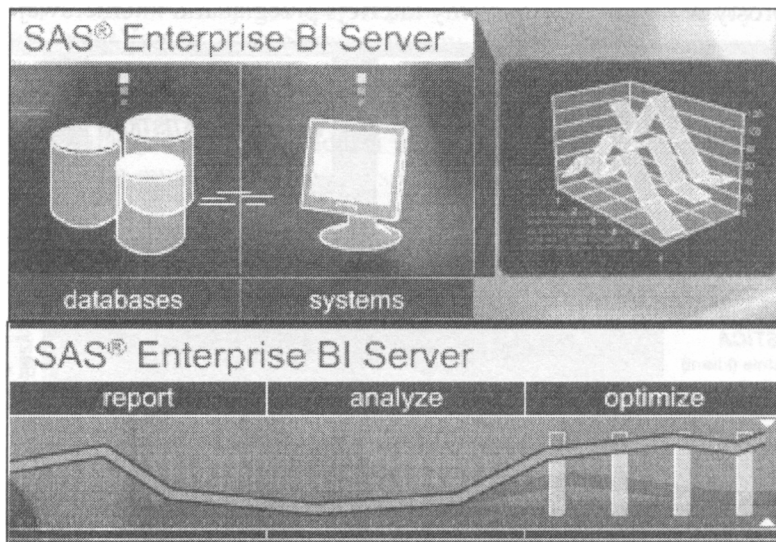
SAS Enterprise Business Intelligence Server³

Rozwiązanie SAS Enterprise BI Server dostarcza zintegrowany i kompleksowy zestaw możliwości oferowanych przez Business Intelligence:

- raportowanie tworzy i udostępnia raporty dla wszystkich poziomów użytkowników w organizacji,
- kwerendy oraz analiza pozwalają na zadawanie pytań i otrzymywanie szybkich odpowiedzi,
- przetwarzanie OLAP analizuje dane *online*,
- zintegrowana analityka łączy narzędzia SAS i tworzy dostępne wyniki,

³Opracowano na podstawie: www.sas.com.

- wizualizacja prezentuje dane w postaci wykresów, schematów i map z różnych aplikacji,
- integracja z Microsoft Office umożliwia dostęp do danych, raportowania i analizy bezpośrednio z Microsoft Office.



Rys. 3. Schemat działania SAS Enterprise Business Intelligence Server

Źródło: www.sas.com.

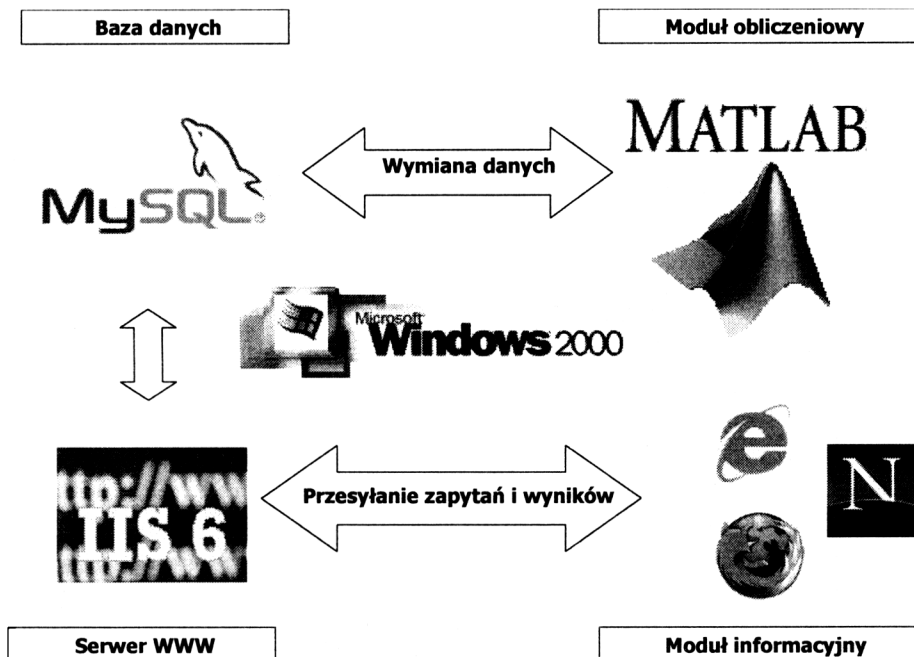
Wykorzystanie dodatkowo SAS Web Report Studio pozwala – z poziomu przeglądarki internetowej – na tworzenie i udostępnianie raportów, wyszukiwanie danych i prowadzenie analiz. Wykorzystanie technologii klient-serwer pozwala na stworzenie prostego narzędzia internetowego umożliwiającego lepszy dostęp do informacji.

3. System Analiz i Prognoz – prezentacja

Technologia SAP oparta jest, podobnie jak przedstawione wyżej przykłady oprogramowania komercyjnego, na internetowej technologii klient-serwer. Pierwotny projekt systemu (zrealizowany jako SPPP) oparty był na platformie operacyjnej Microsoft Windows Server 2003, Matlabie oraz bazie danych Microsoft SQL Server. Obecnie trwają prace nad wykorzystaniem darmowej platformy bazodanowej MySQL. SAP oparto na platformie Microsoft Windows 2000.

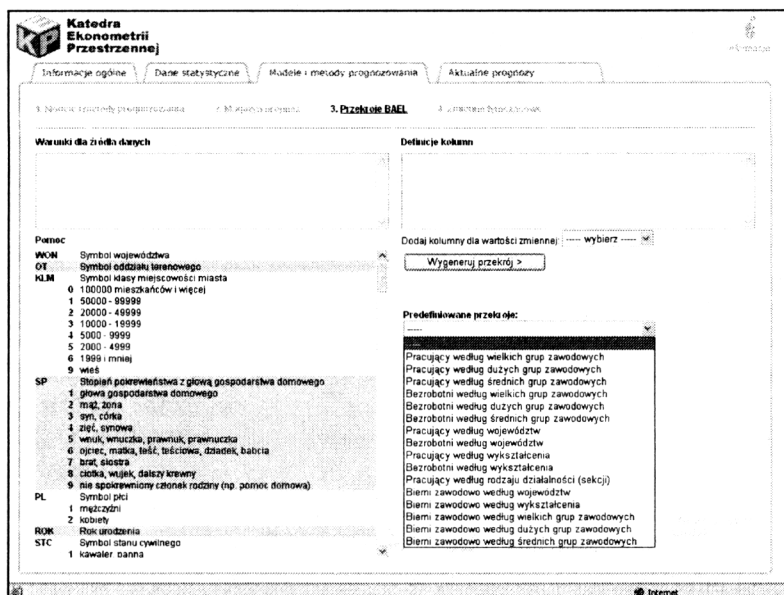
W SAP można wyróżnić następującą strukturę modułową:

- Moduł wejściowy jako bazę danych przechowującą dane w jednolitej formie.
- Moduł obliczeniowy wykonujący analizy i przekazujący ich wyniki do bazy danych.



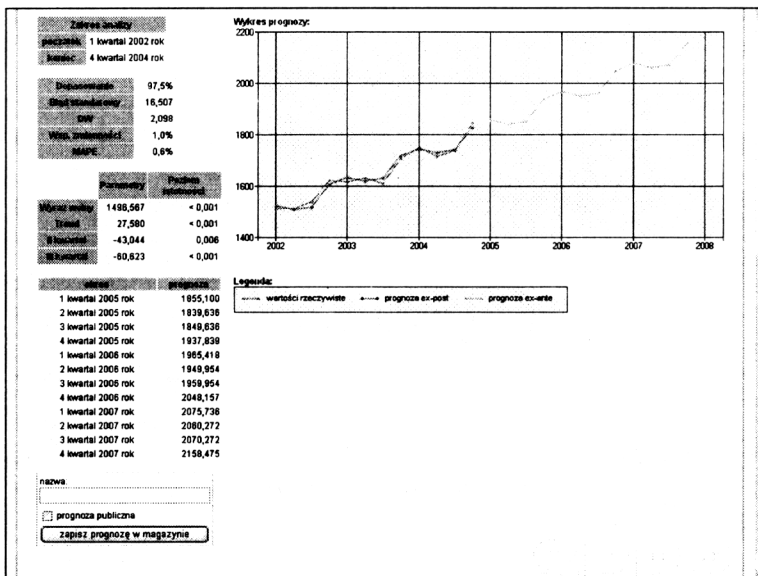
Rys. 4. Schemat działania SAP

Źródło: opracowanie własne na podstawie: www.mathworks.com; www.microsoft.com; www.emuadmin.com; www.netscape.com.

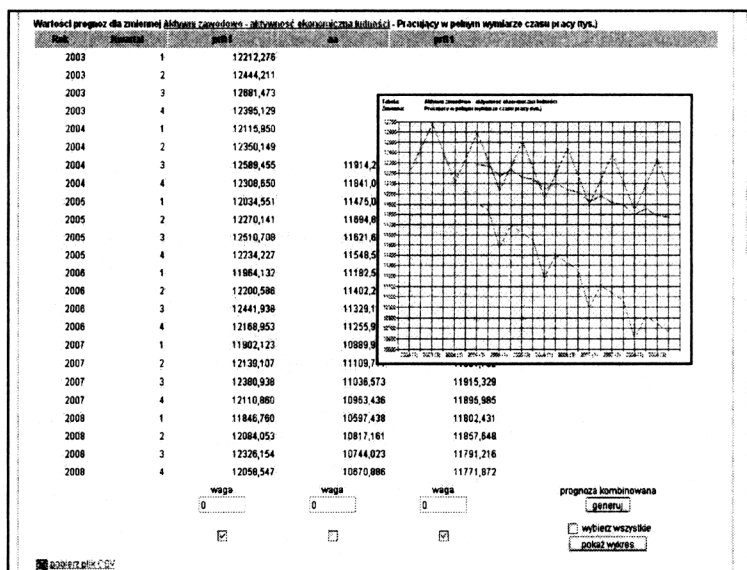


Rys. 5. Baza danych – przekroje BAEL

Źródło: www.sap.uni.lodz.pl.



Rys. 6. Analiza danych
 Źródło: www.sap.uni.lodz.pl.



Rys. 7. Wizualizacja (tabele, wykresy)
 Źródło: www.sap.uni.lodz.pl.

C. Moduł informacyjny i serwer WWW, czyli raporty, tabele, wykresy (prezentacja danych historycznych, prezentacja wyników estymacji i weryfikacji modeli, prezentacja prognoz).

Poza podstawową funkcją magazynowania i prezentowania danych, baza danych SAP pozwala na generowanie dowolnych przekrojów z bazy danych indywidualnych BAEL. Jest to zautomatyzowane źródło unikatowych danych statystycznych.

Moduł obliczeniowy składa się z wyboru modeli analitycznych, magazynu prognoz oraz generatora zmiennych tymczasowych.

Wizualizacja danych i wyników analiz możliwa jest w formie statycznych dokumentów zawierających tekst, tablice i wykresy. Dodatkowo występują interaktywne tablice z danymi liczbowymi, pozwalające na generowanie prognoz kombinowanych i wykresów wybranych szeregów czasowych.

4. Wnioski

SAP stanowi przykład wykorzystania technologii internetowej do budowy systemu wspomagania decyzji, będącego jednocześnie portalem informacyjnym, narzędziem analitycznym oraz hurtownią danych.

Obecnie prace nad systemem dotyczą zmiany platformy bazodanowej i operacyjnej oraz rozbudowy możliwości analitycznych o kolejne modele (np. ARIMA), a także elementy statystycznej analizy i klasyfikacji danych.

Literatura

- Flakiewicz W. (2002), *Systemy informacyjne w zarządzaniu*, C.H. Beck, Warszawa.
- Gajdos A. (2002), *Systemy informacji i prognozowania rynku pracy*, niepublikowana rozprawa doktorska, Łódź.
- Gajdos A., Szablowski P. (2004), *Projekt informatyczny SPPP*, Studia i Materiały t.XII. SPPP, część II. *Projekt informatyczny, baza danych, model makroekonomiczny*, RCSS, Warszawa.
- Microsoft SQL Server 2000 (2001), Kalen Delaney, Wydawnictwo RM, Warszawa.
- Sucheckie B. (2004), *SPPP jako system informacyjny*, Studia i Materiały t. XII. SPPP, część II. *Projekt informatyczny, baza danych, model makroekonomiczny*, RCSS, Warszawa.
- Sucheckie B. (1998), *Koncepcja systemu analiz i prognoz rynku pracy w Polsce*, [w:] Kryńska E., Suchecka B., *Prognoza podaży i popytu na pracę w Polsce do roku 2010*, IPiSS, Warszawa.
- <http://www.mathworks.com/products/matlab/description1.html>.
- <http://www.microsoft.com/poland/sql/evaluation/wyniki.asp>.

DATA ANALYSIS AND FORECASTING SYSTEM – VISUALIZATION AND DATA PROCESSING TOOL

Summary

Data Analysis and Forecasting System, which authors are the team from Chair of Spatial Econometrics, is the on-line information and forecasting system.

System provides connections between databases, analytical tools and new, dynamic WWW technologies. Working in the system rely on calculations on statistical data and presentations results on-line. More information about actual version of the system: www.sap.uni.lodz.pl.